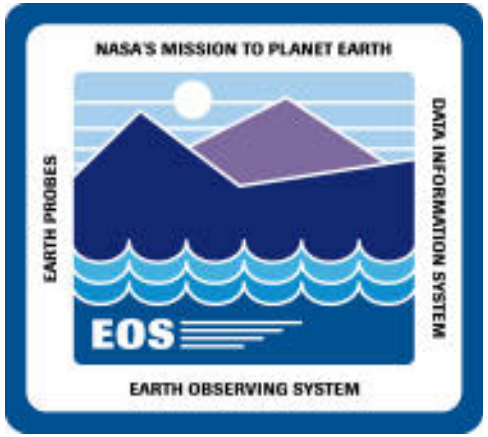


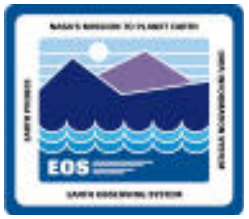
Data Server Subsystem Overview

Parris Caulk

pcaulk@eos.hitc.com

17 April 1996

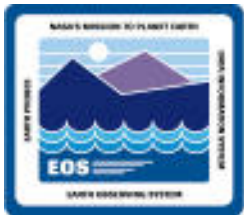
Detailed Design Agenda



17 April 1996

8:00 - 8:45	DSS Status	Paul Fingerman
8:45 - 9:15	DSS Overview	Parris Caulk
9:15 - 10:00	Featured Topics	Michael Burnett
10:00 - 10:15	*Break*	
10:15 - 10:45	Hardware Configuration	Alla Lake
10:45 - 12:00	DDIST Detail Design	Michael Burnett
12:00 - 1:00	Lunch	
1:00 - 2:00	DDIST Detail Design	Michael Burnett
2:00 - 3:00	STMGT Detail Design	Don Brown
3:00 - 3:15	*Break*	
3:15 - 5:00	STMGT Detail Design	Don Brown

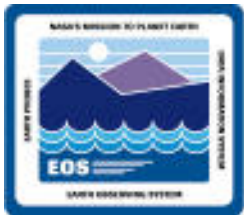
Detailed Design Agenda



18 April 1995

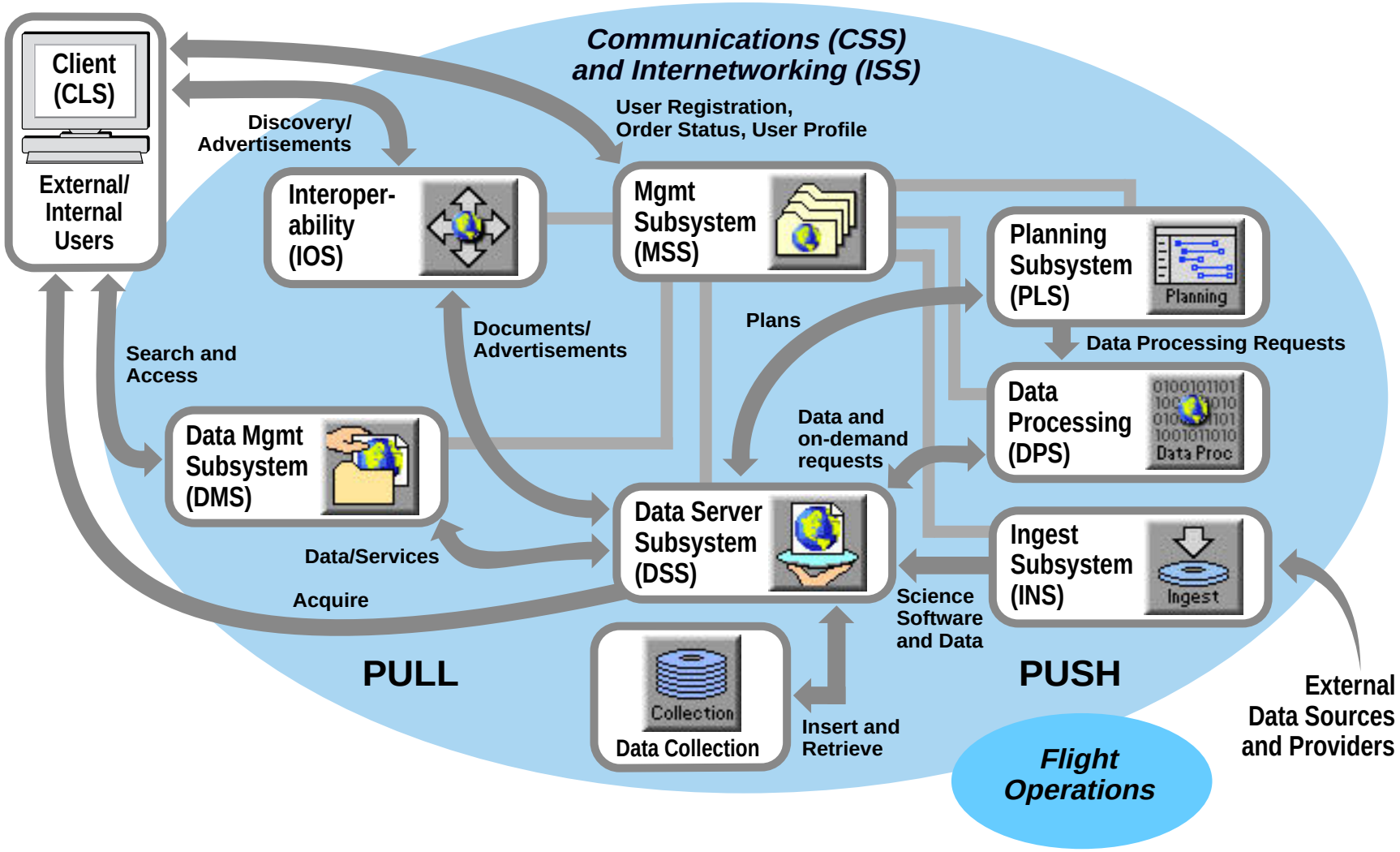
8:00 - 8:45	**SDSRV/DDSRV Public Classes**	**Evelyn Nakamura**
8:45 - 9:15	**Illustra Management Status**	**Jan Dreisbach**
9:15 - 10:00	**SDSRV Design Status**	**Jan Dreisbach**
10:00 - 10:15	***Break***	
10:15 - 11:00	**SDSRV Design Status - cont'd**	**Jan Dreisbach**
11:00 - 11:40	**DDSRV Design Status**	**Evelyn Nakamura**
11:40 - 12:00	**Wrapup**	**Judy Smith**

Data Server Subsystem Overview

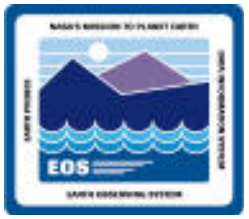


- ECS context
- Driving Requirements
- Software Overview By CSCI
- COTS Selections
- Hardware Overview By CSCI
- Analysis since IDR

ECS Context


Client (CLS)
External/ Internal Users
Communications (CSS) and Internetworking (ISS)
Discovery/ Advertisements
User Registration, Order Status, User Profile
Interoper-ability (IOS)
Mgmt Subsystem (MSS)
Planning Subsystem (PLS)
Planning
Data Processing Requests
Documents/ Advertisements
Plans
Search and Access
Data Mgmt Subsystem (DMS)
Data and on-demand requests
Data Processing (DPS)
Data Proc
Data Server Subsystem (DSS)
Ingest Subsystem (INS)
Ingest
Data/Services
Acquire
Science Software and Data
PULL
PUSH
Collection
Data Collection
Insert and Retrieve
Flight Operations
External Data Sources and Providers

Data Server Subsystem



Manages and provides access to collections of Earth science data

Provides long-term storage for data provided by the Ingest and Data Processing subsystems

Advertises data services to the Interoperability subsystem

Performs services requested by the Client, Data Management, and Data Processing subsystems

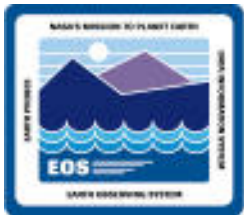
Submits on-demand processing requests to the Data Processing subsystem

Provides temporary storage for the Data Processing subsystem

Stores production plans (documents) for the Planning subsystem

Exchanges system management data with the System Management subsystem

Driving Requirements



Provide general capability to perform services on data, and to add new data types

Provide scalable hardware and software architecture

Maximize the use of COTS software

Provide petabyte storage capacity

Support terabyte average daily I/O throughput

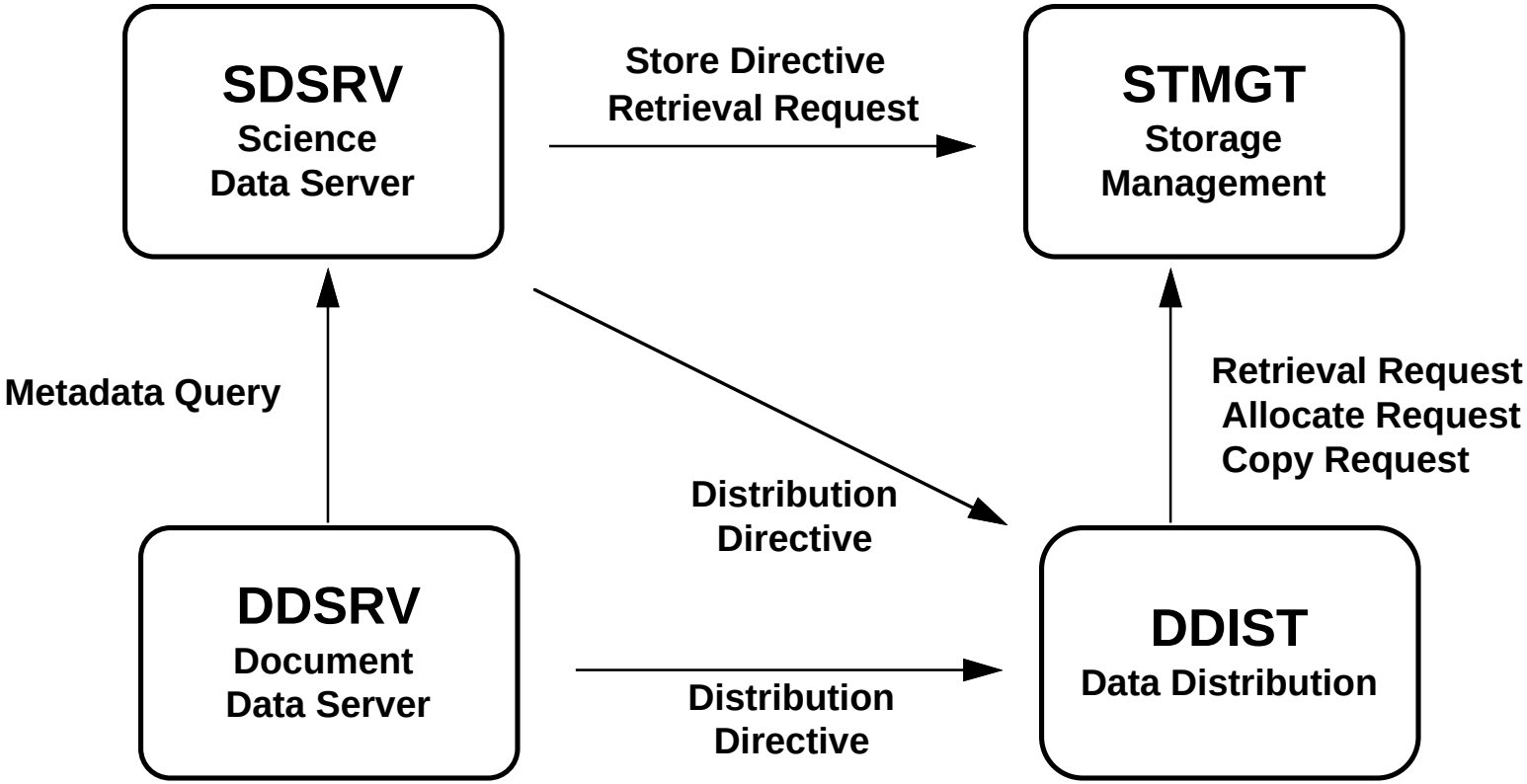
Minimize hardware and operations costs

Enable re-use of components outside of ECS

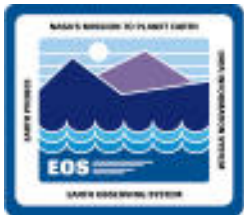
Adhere to standards, “open systems” approach

Maintain data integrity and completeness

Software Components



Software Components (cont.)

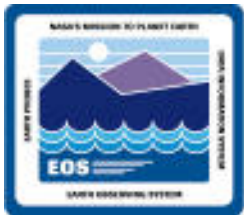


- **SDSRV - Science Data Server**
  - **Includes *Earth Science Data Type* (ESDT) code (products and services)**
  - **Performs data-collection-specific searches**
  - **Manages queues for searches, orders, inserts**
- **STMGT- Storage Management**
  - **Manages <u>all</u> physical storage resource pools for <u>all</u> DSS components**

    Tape robotic archive **Online storage**

    Disk cache **Peripheral devices**
- **DDIST - Data Distribution**
  - **Stages ordered data; services ftp “gets” and “puts”**
  - **Notifies users when ordered data is ready/shipped (on media or electronically)**
- **DDSRV - Document Data Server**
  - **Manages archiving, search, and retrieval of documents**

SDSRV - Science Data Server CSCI



Document Reference 305-CD-024-002 Section 4

Manages logical collections of Earth science data

Advertises data and associated services to external users

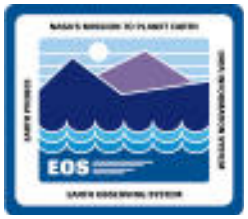
Manages interactive sessions with external clients and processes requests for data access and services

Directs the STMGT and DDIST CSCIs to perform storage and distribution services

Manages storage of metadata

Provides data used as input for data processing

STMGT - Storage Management CSCI



Document Reference 305-CD-024-002 Section 6

Manages high-capacity archive storage hardware and provides associated file access services

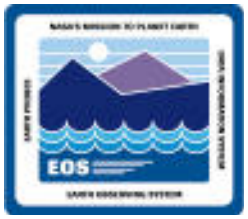
Manages the high-performance storage hardware used for the caching of files in support of the Ingest and Data Processing subsystems

Manages the user-accessable “pull” area used to stage data for electronic distribution

Manages the peripherals used to distribute data to physical media

Manages the storage hardware that composes part of the Ingest subsystem

DDIST - Data Distribution CSCI



Document Reference 305-CD-024-002 Section 7

Formats and distributes data to users

Supports requests from the Science Data Server CSCI

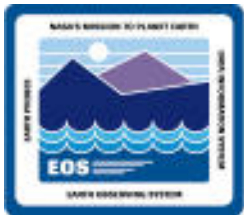
Directs the STMGT CSCI to transfer data to the “user pull” area

Directs the STMGT CSCI to transfer data electronically

Transfers data to 8mm tape, 4mm tape, 6250 bpi 9-track tape, 3480/3490 tape, or CD-ROM

Provides operations the capability to manage distribution requests

DDSRV - Document Data Server CSCI



Document Reference 305-CD-024-002 Section 5

Manages documents and associated metadata

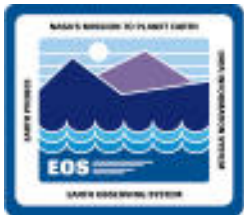
- **Earth science and related documents**
- **Production Plans**

Advertises the availability of documents and associated services to external users

Manages interactive sessions with external clients via the WWW

Provides text search capabilities

COTS Software Selections



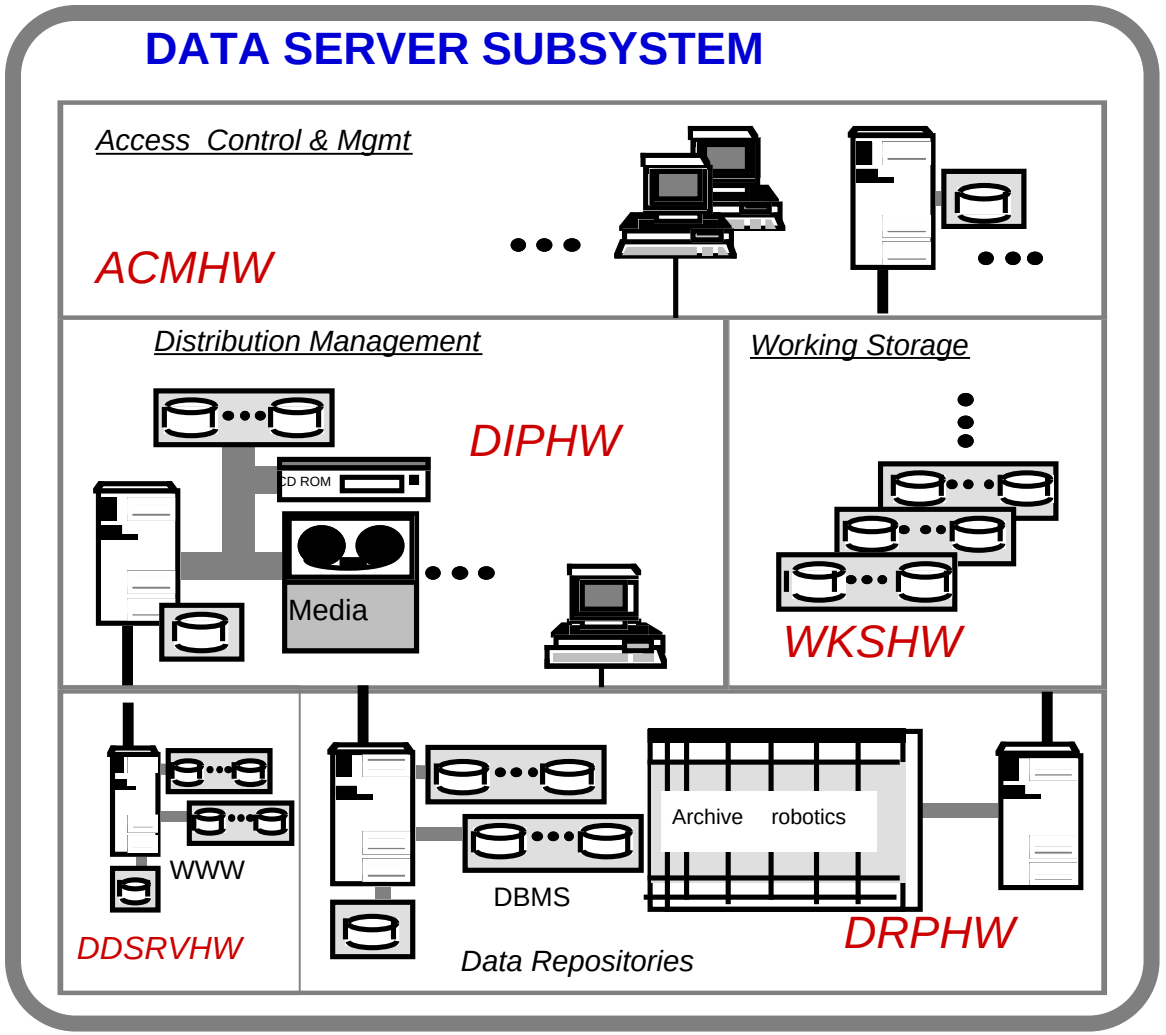
COTS Inherited from Release A

- **Sybase database (STMGT)**
- **Illustra Document Repository (DDSRV)**
- **AMASS file storage management system (STMGT)**
- **Netscape Enterprise Server (DDSRV)**
- **TOPIC Text Search Engine (DDSRV)**

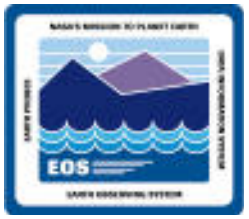
New COTS in Release B

- **Illustra (SDSRV)**

DSS Hardware

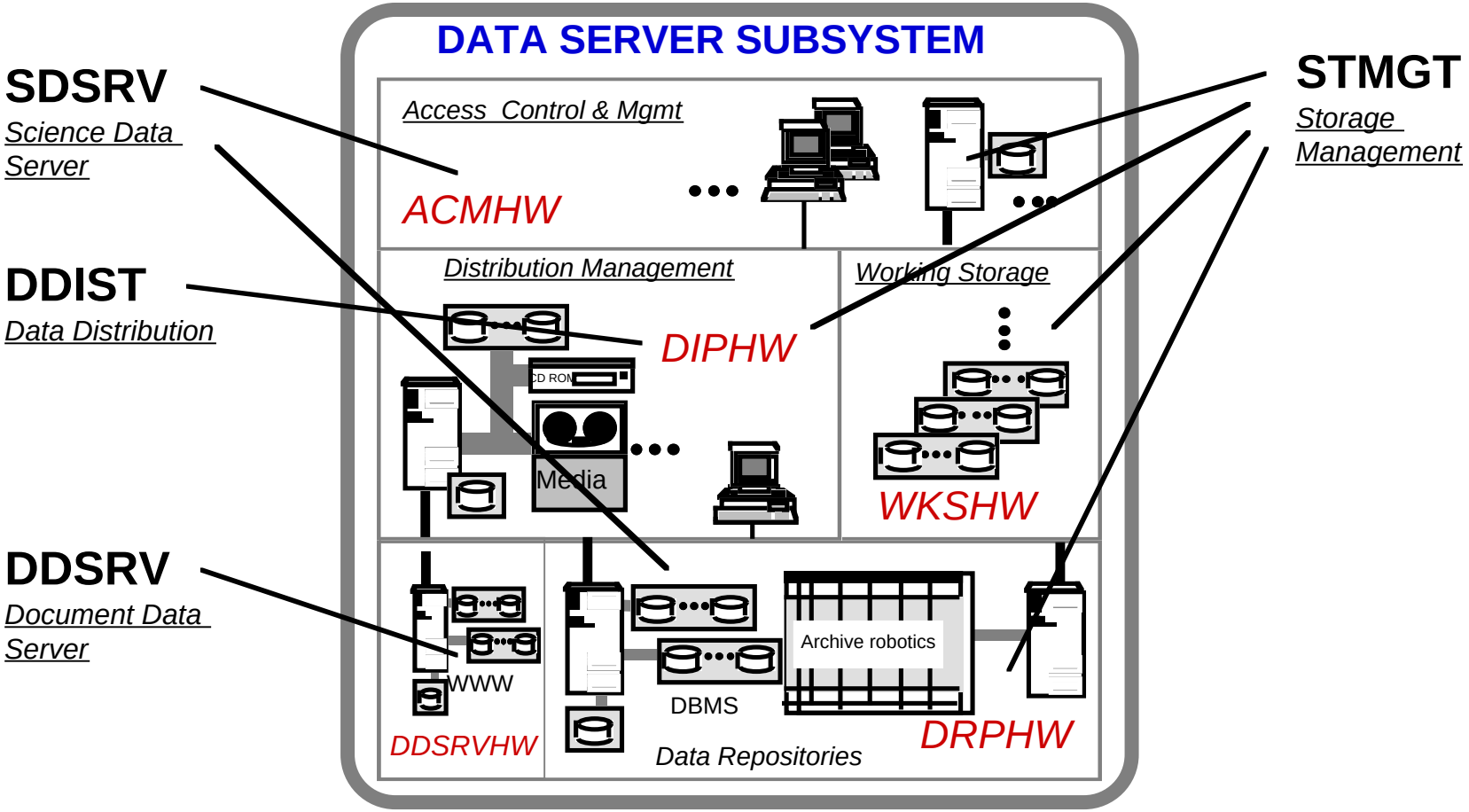


Hardware Configuration Items

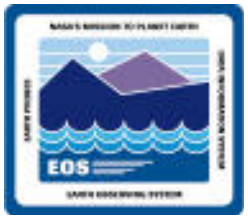


- **ACMHW - Access and Control Management**
  - **Supports the software components of the Ingest and Data Server subsystems that interact directly with users**
- **DRPHW - Data Repository**
  - **Provides high-capacity hardware for the storage, retrieval, and maintenance of data files**
- **DIPHW - Distribution and Ingest Peripherals**
  - **Provides the peripherals and associated processors required for ingest and distribution via physical media**
- **WKSHW - Working Storage**
  - **Provides high-performance storage for the purpose of caching large volumes of data on a temporary basis**
- **DDSRVHW - Document Data Server**
  - **Provides the hardware necessary to store and access documents**

Hardware/Software Mapping



ACMHW- Access and Control Management HWCI



Document Reference 305-CD-024-002 Section 8

Supports the DSS components that interact directly with the users

Hosts the following Data Server subsystem software:

- **SDSRV CSCI - manages client sessions and the processing of service requests**
- **STMGT CSCI - management of user pull area, including a supporting Sybase database**

Hosts the following Ingest subsystem software:

- **Software for non-Level 0 ingest via external networks**

DRPHW- Data Repository HWCI

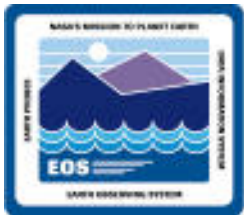
Document Reference 305-CD-024-002 Section 10

Provides high-capacity storage for long-term storage of data files

Hosts the following Data Server subsystem software:

- **SDSRV CSCI - Illustra database engine and supporting software**
- **STMGT CSCI - AMASS file storage management system and interface wrapper**

DIPHW- Distribution and Ingest Peripherals HWCI



Document Reference 305-CD-024-002 Section 11

Provides hardware necessary to support ingest and distribution via physical media

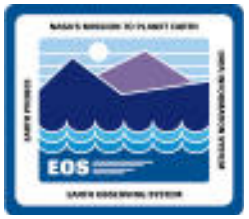
Hosts the following Data Server subsystem software:

- **DDIST CSCI - software for media distribution**
- **STMGT CSCI - software for peripheral management**

Hosts the following Ingest subsystem software:

- **Software for media ingest**

WKSHW- Working Storage HWCI



Document Reference 305-CD-024-002 Section 9

Managed by STMGT software

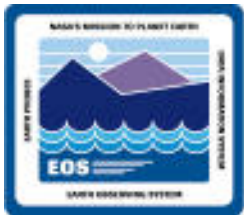
Provides Storage used by the Data Server subsystem to:

- **Cache intermediate data products generated by the Data Processing subsystem**
- **Stage data for processing by the Data Processing subsystem**
- **Stage ingested data until it can be stored in the long-term storage facilities of the DRPHW CSCI**

Provides Storage used by the Ingest subsystem to:

- **Prepare L0 data prior to archiving by the Data Server subsystem**

DDSRVHW- Document Data Server HWCI



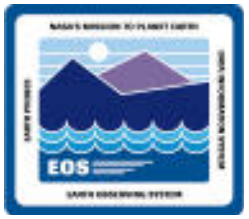
Document Reference 305-CD-024-002 Section 12

Provides the hardware necessary to store, search, and access documents

Hosts the following DDSRV CSCI software

- **Illustra Document Repository**
- **Netscape Server**
- **TOPIC Text Search Engine**
- **Custom software for processing document service requests**

Design Analysis



Data rate analysis for bulk data flows used for hardware sizing

Data Repository (DRPHW) load sensitivity model

- **Results showed required number of robot arms and drives for various loads**
- **Used as input to ECS dynamic model**

Analytical model provided CPU and memory sizing information

Document Storage Trade (DRPHW vs. DDSRVHW)